AWARD NUMBER: W81XWH-14-1-0107

TITLE: Tumor Microenvironment Gene Signature as a Prognostic Classifier and Therapeutic Target

PRINCIPAL INVESTIGATOR: Sandra Orsulic, PhD

CONTRACTING ORGANIZATION: Cedars-Sinai Medical Center
Los Angeles, CA 90048

REPORT DATE: June 2016

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
June 2016	Annual	5May2015 - 4May2016

4. TITLE AND SUBTITLE
Tumor Microenvironment Gene Signature as a

Prognostic Classifier and Therapeutic Target

5a. CONTRACT NUMBER
W81XWH-14-1-0107

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Sandra Orsulic

E-Mail: orsulics@cshs.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cedars-Sinai Medical Center
8700 Beverly Boulevard
Los Angeles, CA 90048

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We identified a tumor microenvironment-based activated fibroblast gene signature that correlates with poor survival in ovarian cancer patients. We are refining this gene signature to develop biomarkers for the identification of patients with adverse outcomes on standard treatment. In the first part of this project, we have analyzed a gene signature for the identification of patients who are unlikely to benefit from standard surgery and/or chemotherapy and should be considered for clinical trials targeting specific pathways in the tumor microenvironment. Specifically, we found that suboptimal surgical outcome is associated with a molecularly aggressive subtype of ovarian cancer characterized by the presence of activated fibroblasts, which likely contributes to chemotherapy resistance. In the current part of the project, we focused on the molecular characterization of activated cancer fibroblasts. First, we defined COL11A1 as a highly specific biomarker of *activated* fibroblasts. Second, using COL11A1 as a 'seed' to identify coexpressed genes, we demonstrated that activated fibroblasts express a highly conserved gene signature across genetically different epithelial cancer types. In the last part of the project (no-cost extension), we will validate the gene signature in patient samples and develop a preliminary quantitative assay for use in the clinical setting.

15. SUBJECT TERMS
Ovarian cancer, tumor microenvironment, gene signature, carcinoma-associated fibroblasts, clinical outcome

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	19	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. INTRODUCTION:

The majority of patients with advanced stage epithelial ovarian cancer (EOC) present with advanced stage disease, which is currently treated by cytoreductive surgery and chemotherapy. Approximately 10% of EOC patients cannot be successfully cytoreduced by surgery and 20% are intrinsically resistant to chemotherapy or develop chemoresistant disease within one year from initial treatment. Currently, ovarian cancer surveillance and subsequent therapies are implemented on a "watch-and-wait" basis because there are no reliable biomarkers to identify patients with adverse outcomes on standard treatment. To identify biomarkers that predict adverse outcome in patients, we studied the key processes involved in metastatic ovarian cancer progression, including changes in the tumor microenvironment. This led to the identification of a stromal/extracellular matrix gene signature that correlates with poor patient survival. In the first part of this project, we have identified and optimized gene signatures for the identification of patients who are unlikely to benefit from standard surgery and/or chemotherapy. Since most of the signature genes were expressed in the cancer stroma (fibroblasts and extracellular matrix secreted by fibroblasts), we focused on the molecular characterization of cancer fibroblasts. Cancer fibroblasts actively contribute to ovarian cancer progression and are viewed as a promising therapeutic target. However, the role of cancer fibroblasts as collaborators in cancer progression has recently been challenged. Ablation of α-smooth muscle actin (αSMA)-positive fibroblasts in a mouse model of pancreatic cancer has resulted in increased cancer progression and metastasis, leading authors to conclude that fibroblasts have a preventive role in the progression of some cancers [1]. Another recent study in a mouse model of pancreatic cancer concluded that cancer genotype determines the activation state of cancer fibroblasts [2].

The contradictory results in different cancer models could be explained by different roles of fibroblasts in different cancer types, i.e. fibroblasts could be promoting ovarian cancer and inhibiting pancreatic cancer. Alternatively, in all cancer types fibroblasts prevent cancer progression until they receive activating signals from cancer cells and convert into 'activated fibroblasts', which in turn confer invasive and metastatic abilities upon cancer cells [3]. Therapies that target all fibroblasts are counterproductive and likely to result in the death of normal fibroblasts and significant toxicity.

Preferential targeting of activated fibroblasts has been challenging because activated fibroblasts are poorly understood at the molecular level. During activation, fibroblasts exhibit phenotypic changes that partially overlap with myofibroblastic changes during wound healing, inflammation, and fibrosis, including secretion of specific ECM components, cytokines and growth factors [4, 5]. Several markers have been used to distinguish activated from non-activated fibroblasts: α-smooth muscle actin (αSMA, encoded by gene *ACTA2*), fibroblast activation protein (FAP), podoplanin (PDPN), palladin (PALLD), tenascin-C (TNC), platelet-derived growth factor receptor α (PDGFRα), and NG2 chondroitin sulfate proteoglycan (CSPG4). However, these markers are frequently expressed in other cells within the cancer stroma, such as vascular smooth muscle cells, pericytes, and mesenchymal stem cells. This lack of specificity could pose problems in therapeutic targeting and underscores the need to better understand the molecular characteristics of activated fibroblasts in order to develop more precise and less toxic targeted therapies.

We previously identified *COL11A1* as part of gene signatures associated with suboptimal debulking, rapid recurrence and poor overall survival in ovarian cancer [6, 7]. *In situ* hybridization and immunohistochemistry in ovarian cancer revealed that *COL11A1* mRNA and pro-protein are primarily expressed in cancer fibroblasts [7, 8]. We explored the suitability of *COL11A1* as a pan-cancer marker of *activated* fibroblasts and used it as a 'seed' to identify the transcription signature of activated fibroblasts in 13 epithelial cancer types. We show that the *COL11A1*-coexpressed gene set is highly conserved in these 13 cancer types, indicating that the fibroblast reaction to cancer cells is independent of the organ site-of-origin and of the transforming events within cancer cells.

2. KEYWORDS:

Ovarian cancer, tumor microenvironment, gene signature, carcinoma-associated fibroblasts, clinical outcome

3. ACCOMPLISHMENTS:

- **What were the major goals of the project?**

Specific Aim 1 (specified in proposal)	Timeline	% Completed
Major Task 1 **Identify gene signatures for the prediction of poor outcome**	**Months**	
Subtask 1 Select biologically relevant covariates and build a multivariate model for the analysis of 3 datasets (TCGA, n=403; GSE26712, n=185; and GSE51088, n=122; these are public datasets with de-identified patient information).	1-3	100%
Subtask 2 Analyze datasets individually and in combination; derive gene signatures using multiple statistical methods; correlate with overall survival, progression-free survival, residual disease and other outcomes	4-6	100%
Subtask 3 Identify small subsets of predictive genes and their interactions	7-8	100%
Milestone Achieved A gene signature consisting of 8-15 genes with high predictive power in all three datasets	8	100%
Major Task 2 **Optimize the gene signature for the prediction of poor outcome**		
Subtask 1 Assess the predictive accuracy of the gene signatures using independent datasets (GSE9891 and GSE3149)	9-10	100%

Subtask 2 Generate a test qPCR set from frozen samples of patients with extreme outcomes (10 with <1 year survival and 10 with >7 year survival); validate by qPCR up to 30 prioritized genes that have extremely high predictive power in individual datasets but are not present in all 3 datasets	11-12	10%
Subtask 3 Validate gene signature accuracy using statistical methods	13-14	0%
Milestone Achieved An independently validated set of 8-15 genes with high predictive power	14	0%
Major Task 3 **Validate the gene signature for the prediction of poor outcome**		
Subtask 1 Identify and collect 200 primary ovarian cancer patient samples with annotated demographic, pathologic and clinical information and follow-up (all patient samples will be de-identified)	3-7	100%
Subtask 2 Cut and stain slides (1 H&E + 9 unstained sections); evaluate the suitability of each sample by pathologic examination of tumor sections and circle the area on the slide for RNA isolation	8-9	100%
Subtask 3 Isolate RNA from slides; perform quality control	9-12	100%
Subtask 4 Design the Nanostring assay for the quantification of the signature genes; perform quality control; collect data	13-15	100%
Subtask 5 Analyze data using statistical methods; correlate with overall survival, progression-free survival, residual disease status and other outcomes that could be used to improve clinical management of ovarian cancer patients	16-21	0%
Subtask 6 Validate the gene signature assay using statistical models and risk prediction models with known parameters	21-22	0%
Subtask 7 Submit a manuscript on the predictive power of the optimized gene signature using microarray data and Nanostring assay data and deposit RNA expression data into public repository (GEO) Plan an academic multi-center validation of the Nanostring signature assay as required prior to FDA validation	23-24	0%
Milestone Achieved Validated gene signature gene assay for the prediction of clinical outcome(s) using paraffin-embedded tumor tissues	*22*	0%

- **What was accomplished under these goals?**

1) major activities

In the first part of this project, we have identified gene signatures for the identification of patients who are unlikely to benefit from standard surgery and/or chemotherapy. Since most of the signature genes were expressed in the cancer stroma (fibroblasts and extracellular matrix secreted by fibroblasts), we focused on the molecular characterization of cancer fibroblasts.

2) specific objectives

To effectively target fibroblasts in ovarian cancer, it is necessary to accurately characterize the molecular characteristics of the activated fibroblast populations in different cancer types and determine whether ovarian cancer activated fibroblasts are different from activated fibroblasts in other cancer types. Considering the artificial nature and limitations of mouse cancer models, it is important to conduct this analysis in human tissues.

Our objective was to conduct an in-depth analysis of public expression datasets. First, we defined *COL11A1* as a highly specific biomarker of *activated* fibroblasts in *multiple* epithelial cancer types. Second, using COL11A1 as a 'seed' to identify coexpressed genes, we demonstrated that activated fibroblasts express a highly conserved gene signature across genetically different epithelial cancer types.

3) significant results or key outcomes, including major findings, developments, or conclusions

3.3.1. COL11A1 is associated with cancer progression and adverse clinical outcomes in ovarian cancer

COL11A1 mRNA expression has been associated with poor survival in ovarian cancer [7, 9]. To elucidate the underlying biology that could result in poor survival, we investigated its expression in ovarian and colon cancers. Using a comprehensively annotated microarray database for 3431 human ovarian cancers [10], we show that increased expression of *COL11A1* mRNA is associated with disease-specific and disease-free survival (**Fig. 1A**) as well as with clinical and molecular parameters such as increased cancer stage and grade and mesenchymal molecular subtype (**Fig. 1B**).

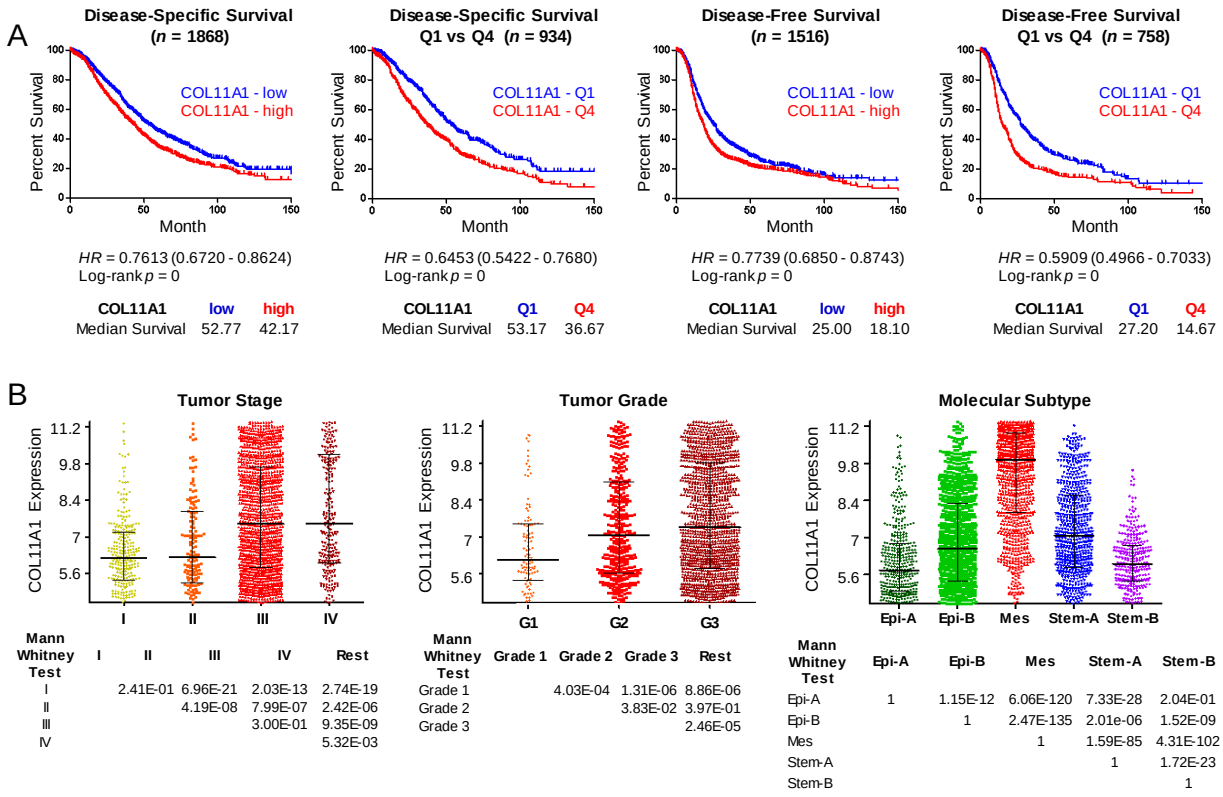

Fig. 1. ***COL11A1*** **expression is associated with adverse clinical parameters. (A)** Kaplan-Meier survival plots and **(B)** plots of *COL11A1* expression in individual subsets of ovarian cancer were generated using the ovarian microarray gene expression database CSIOVDB (csibio.nus.edu.sg/CSIOVDB/CSIOVDB.html).

3.3.2. COL11A1 is expressed in a subset of αSMA-positive cancer fibroblasts and can be induced in normal fibroblasts by the presence of cancer cells

To determine if *COL11A1* expression is associated with fibroblast activation, we used αSMA as a marker of activated fibroblasts [11]. Comparison of αSMA immunohistochemistry and *COL11A1 in situ* hybridization in a tissue microarray consisting of primary, metastatic and recurrent ovarian cancers from 42 patients showed that *COL11A1* is expressed in a subset of αSMA+ fibroblasts (**Fig. 2A**). Unlike αSMA, *COL11A1* was not expressed in blood vessels (red arrows) or in fibroblasts surrounding the cancer (blue arrows) (**Fig. 2A**). In sections of metastatic ovarian cancer, we observed that *COL11A1*-positive cells are confined to the intratumoral and immediate peritumoral fibroblasts (**Fig. 2B**), suggesting that *COL11A1* expression may be induced by cues received from epithelial cancer cells. To test if cancer cells can induce *COL11A1* expression in fibroblasts, we co-cultured immortalized normal ovarian fibroblasts (INOFs) with three different ovarian cancer cell lines (OVSAHO, OVCAR3, and KURAMOCHI). *COL11A1* expression in INOFs was most strongly induced by direct co-culture with ovarian cancer cell lines although weak induction occurred by indirect co-culture on a Transwell membrane (**Fig. 2C**). The induction of *COL11A1* in fibroblasts in the presence of cancer cells was confirmed by analysis of the public expression dataset GSE52104 in which two

types of presumptive cancer-associated fibroblast precursor cells, mesenchymal stem cells (MSCs) and immortalized normal ovarian fibroblasts (INOFs), were either cultured alone or co-cultured with normal ovarian surface epithelial cells (IOSE) or epithelial ovarian cancer cells (EOC) using a Transwell membrane [12]. *COL11A1* mRNA was statistically significantly upregulated when MSCs and INOFs were co-cultured with EOC but not IOSE (**Fig. 2D**), indicating that cancer cells have a greater capacity than normal cells to induce *COL11A1* expression in fibroblasts.

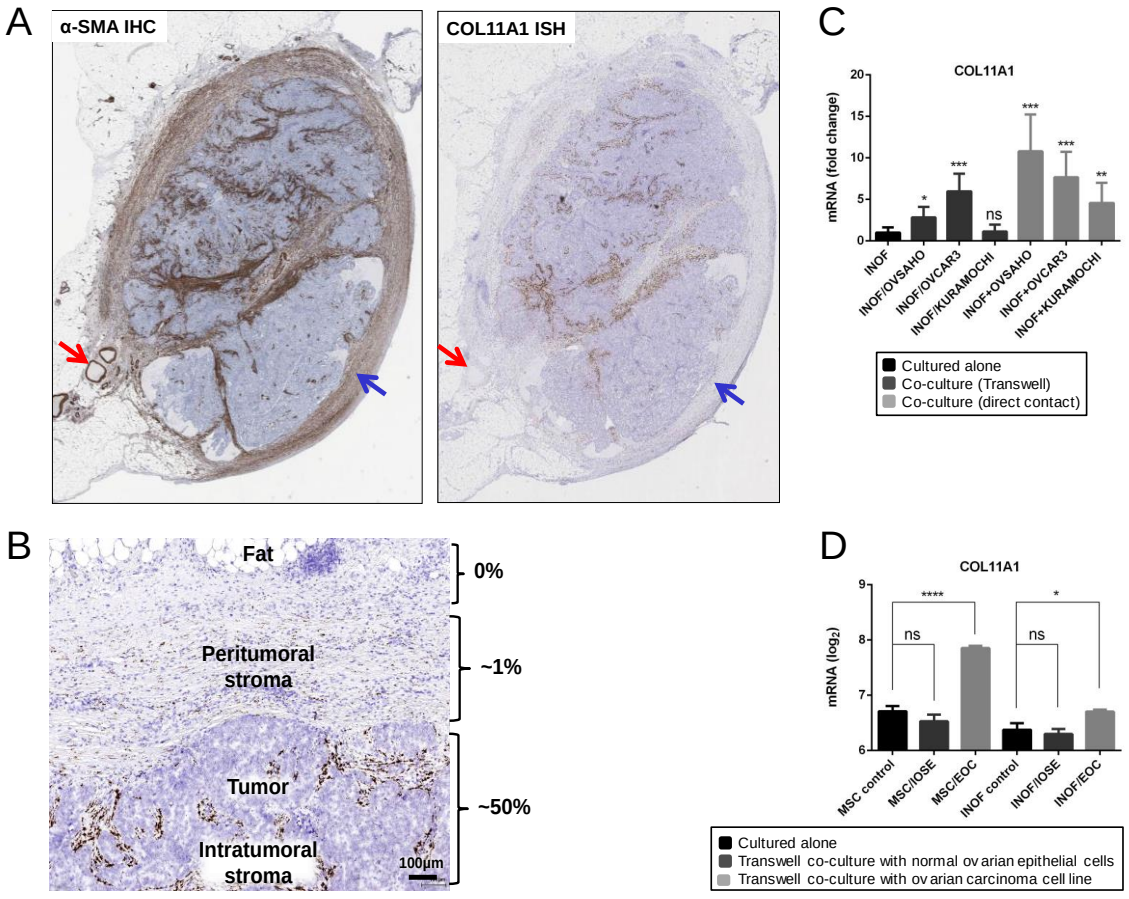

Fig. 2. ***COL11A1* is expressed in cancer-activated fibroblasts. (A)** Comparison of αSMA immunohistochemistry and *COL11A1 in situ* hybridization in a metastatic ovarian cancer sample. Red arrows indicate blood vessels. Blue arrows indicate fibroblasts surrounding the tumor nodule. **(B)** Distribution of *COL11A1*-positive fibroblasts in relation to cancer cells. The estimated percent of *COL11A1*-positive fibroblasts is shown on the right. The image is representative of metastatic and recurrent ovarian cancer samples, which typically express higher levels of *COL11A1* than primary ovarian cancers. **(C)** Quantitative RT-PCR levels of *COL11A1* in sorted (FACS) immortalized normal ovarian fibroblasts (INOFs) grown alone or co-cultured with ovarian cancer cell lines (OVSAHO, OVCAR3, KURAMOCHI) that were either separated from INOFs by a Transwell membrane or directly mixed with INOFs. Statistical analyses were performed between INOFs grown alone and INOFs co-cultured with ovarian cancer cells (*p<0.05; **p<0.01; ***p<0.001; ns, not significant). Error bars indicate standard deviation. **(D)** Levels of *COL11A1* in the GSE52104 expression dataset in which mesenchymal stem cells

(MSCs) or immortalized normal ovarian fibroblasts (INOFs) were either cultured alone or co-cultured with IOSE4 normal epithelial cells (IOSE) or HEYA8 epithelial ovarian cancer cells (EOC) using a Transwell membrane. Inverse-log2 data of different *COL11A1* probes were averaged, then log2-transformed. The data were extracted for statistical analyses using GraphPad Prism 6. Data are represented as the mean ± SEM. Intergroup differences were assessed by the Student's t-test. *p<0.05; ****p<0.0001.

3.3.3. COL11A1 is among the most differentially expressed genes between cancers and corresponding benign tissues

Comparison of *COL11A1* expression in 17931 cancers and 3503 normal tissues (U133Plus2 platform) and 9258 cancers and 4087 normal tissues (U133A platform) using the Gene Expression across Normal and Tumor tissue (GENT) portal [13] revealed that *COL11A1* mRNA is elevated in most cancers in comparison to their corresponding normal tissues (**Fig. 3A)**. As many collagens and collagen-remodeling genes are frequently upregulated in fibroblast activation associated with inflammation and fibrosis in the absence of cancer, use of these genes as therapeutic targets in cancer could be problematic. Analysis of expression profile datasets show that levels of *COL11A1* mRNA in inflamed colonic tissue from inflammatory bowel disease and fibrotic lung tissues are not significantly different from those in corresponding unaffected colon and lung and that *COL11A1* levels associated with colon inflammation and lung fibrosis are minimal and markedly different from those associated with colon and lung cancers (**Fig. 3B**). In contrast, levels of *ACTA2*, the gene encoding the prototypical marker of myofibroblast differentiation, αSMA, is expressed at similar levels in cancers and inflamed or fibrotic tissues (**Fig. 3B**).

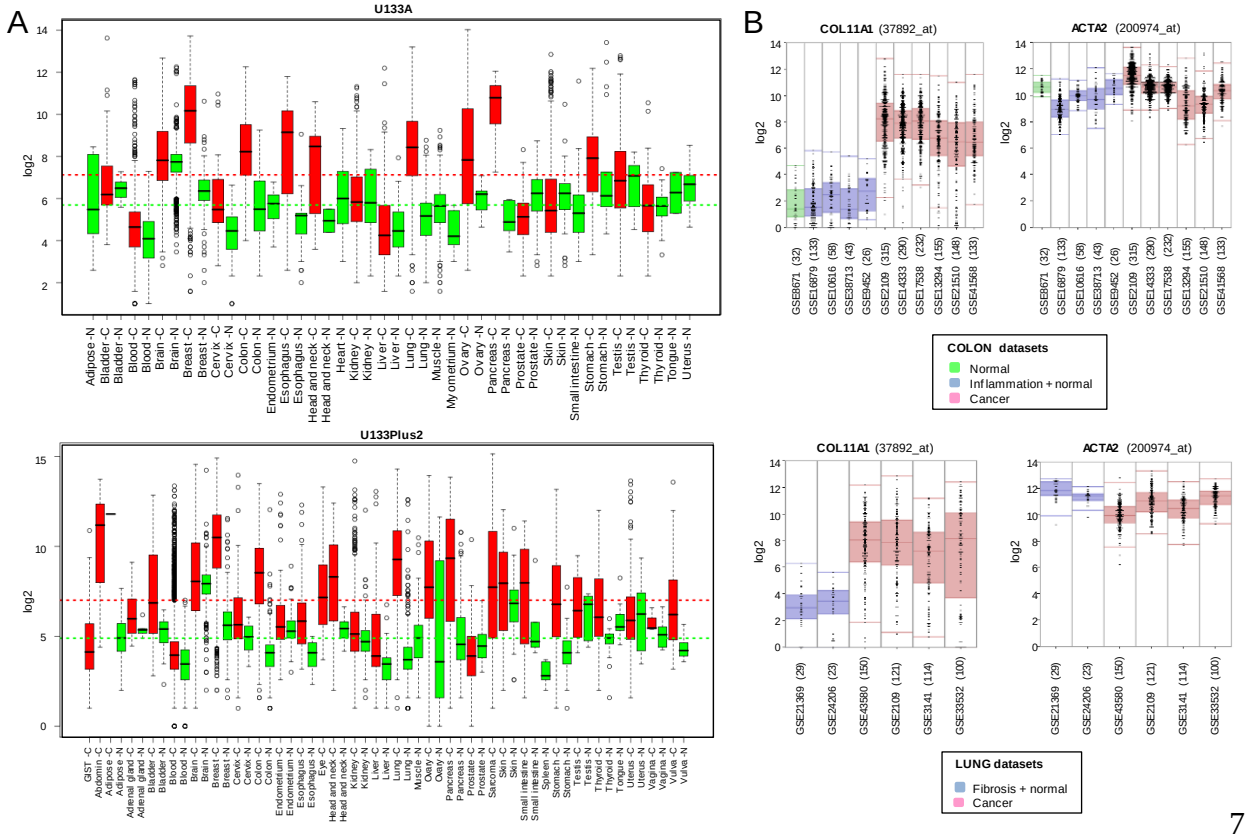

Fig. 3. Increased expression of *COL11A1* in cancer and low expression in normal tissues, inflammation and fibrosis. (A) Comparison of *COL11A1* mRNA expression in normal tissues and corresponding cancers. Box plots for two different platforms (U133Plus2 and U133A) were generated using datasets and software available through the Gene Expression across Normal and Tumor tissue (GENT) portal (medical-genome.kribb.re.kr/GENT). The y axis shows log2 mRNA levels. Average expression levels in normal tissues and cancer tissues are indicated by vertical dotted green and red lines, respectively. **(B)** *COL11A1* and ACTA2 mRNA expression in normal, inflammatory and fibrotic conditions in comparison to cancer. The graphs were generated using the public R2 MegaSampler software (hgserver1.amc.nl/cgi-bin/r2/main.cgi) for the processing and normalization of individual datasets imported from the Gene Expression Omnibus (u133p2, MAS5.0 platform). The number of samples in each GSE dataset are indicated in parentheses.

3.3.4. A consistent set of genes is co-expressed with *COL11A1* across different cancers

To better understand the biology of cancers with high levels of *COL11A1*, we identified genes that most closely correlate with *COL11A1* mRNA expression in 13 TCGA datasets representing different cancer types. Spearman's rank correlations between *COL11A1* and its co-expressed genes for each cancer type were calculated. The genes were then ranked based on the average correlation of each gene across the 13 cancer types. The top 195 correlated genes were selected based on an average correlation of > 0.4 *COL11A1*-correlated genes were then ranked based on the average of the absolute correlation values (**Table 1**). The top 10% most highly correlated genes in each cancer type are highlighted in pink (**Table 1**). Notably, *COL11A1*-correlated genes with high average correlation scores also tended to be among the top 10% highest scored genes in each cancer type (indicated in pink in **Table 1**).

Pan-carcinoma COL11A1-correlated genes

	Gene Symbol	Bladder	Breast	Colorectal	Cervical	Head and Neck	Kidney Clear Cell	Kidney Papillary	Lung Adeno	Lung Squamous	Ovarian	Prostate	Stomach	Thyroid	Average Correlation
1	COL11A1														
2	COL1A1	0.87	0.69	0.76	0.72	0.87	0.72	0.67	0.8	0.86	0.83	0.63	0.74	0.85	0.77
3	COL1A2	0.85	0.74	0.86	0.71	0.87	0.69	0.58	0.8	0.88	0.8	0.51	0.73	0.75	0.752
4	COL3A1	0.85	0.7	0.88	0.64	0.85	0.68	0.63	0.83	0.86	0.85	0.54	0.66	0.79	0.751
5	FAP	0.77	0.72	0.83	0.55	0.69	0.75	0.63	0.77	0.77	0.85	0.59	0.73	0.84	0.73
6	COL5A1	0.83	0.75	0.82	0.59	0.74	0.71	0.65	0.75	0.82	0.82	0.45	0.66	0.75	0.718
7	CTHRC1	0.8	0.68	0.87	0.37	0.81	0.55	0.54	0.81	0.74	0.69	0.61	0.77	0.8	0.695
8	SULF1	0.85	0.76	0.88	0.63	0.79	0.57	0.54	0.84	0.75	0.62	0.47	0.76	0.55	0.693
9	VCAN	0.76	0.72	0.87	0.67	0.83	0.42	0.2	0.79	0.81	0.83	0.54	0.62	0.84	0.685
10	FN1	0.74	0.84	0.69	0.71	0.78	0.56	0.42	0.66	0.78	0.82	0.43	0.61	0.66	0.669
11	SFRP2	0.82	0.54	0.73	0.57	0.79	0.68	0.54	0.81	0.79	0.81	0.32	0.46	0.84	0.669
12	OLFML2B	0.78	0.65	0.7	0.59	0.79	0.56	0.57	0.74	0.71	0.77	0.48	0.56	0.66	0.658
13	COL6A3	0.79	0.66	0.82	0.56	0.78	0.69	0.58	0.66	0.81	0.76	0.23	0.52	0.7	0.658
14	INHBA	0.63	0.78	0.76	0.53	0.37	0.57	0.53	0.77	0.73	0.87	0.52	0.75	0.73	0.657
15	ASPN	0.82	0.71	0.79	0.4	0.75	0.54	0.49	0.71	0.77	0.8	0.61	0.46	0.68	0.656
16	ADAMTS12	0.76	0.69	0.8	0.61	0.73	0.52	0.47	0.69	0.85	0.8	0.28	0.73	0.53	0.651
17	LUM	0.73	0.76	0.81	0.46	0.68	0.7	0.51	0.65	0.65	0.85	0.26	0.52	0.84	0.648
18	NTM	0.78	0.74	0.88	0.68	0.8	0.5	0.51	0.58	0.73	0.84	0.4	0.69	0.27	0.646
19	SPARC	0.77	0.69	0.84	0.66	0.83	0.42	0.25	0.73	0.84	0.81	0.51	0.68	0.34	0.644
20	AEBP1	0.81	0.71	0.76	0.54	0.81	0.46	0.39	0.68	0.85	0.78	0.39	0.48	0.66	0.64
21	CDH11	0.74	0.74	0.76	0.52	0.73	0.53	0.55	0.51	0.75	0.7	0.49	0.55	0.75	0.64
22	GREM1	0.74	0.62	0.74	0.56	0.67	0.58	0.54	0.81	0.79	0.79	0.46	0.27	0.71	0.637
23	ITGBL1	0.67	0.64	0.71	0.63	0.65	0.55	0.6	0.51	0.67	0.77	0.69	0.42	0.76	0.636
24	FNDC1	0.87	0.72	0.79	0.55	0.72	0.61	0.49	0.74	0.89	0.61	0.61	0.63	0.03	0.635
25	PRRX1	0.79	0.67	0.78	0.47	0.68	0.63	0.55	0.76	0.69	0.79	0.17	0.58	0.61	0.628
26	MMP11	0.7	0.71	0.58	0.51	0.78	0.48	0.09	0.71	0.69	0.81	0.51	0.76	0.81	0.626
27	FBN1	0.74	0.75	0.88	0.62	0.72	0.58	0.37	0.67	0.85	0.82	0.35	0.45	0.26	0.62
28	CTSK	0.78	0.65	0.77	0.46	0.78	0.55	0.11	0.75	0.8	0.81	0.35	0.51	0.69	0.616
29	COL8A1	0.65	0.75	0.83	0.65	0.7	0.38	0.43	0.52	0.79	0.77	0.46	0.45	0.48	0.605
30	MMP2	0.73	0.59	0.79	0.51	0.68	0.61	0.42	0.65	0.83	0.8	0.22	0.41	0.62	0.605
31	WISP1	0.76	0.66	0.72	0.45	0.63	0.4	0.48	0.73	0.71	0.71	0.24	0.58	0.77	0.603
32	COMP	0.66	0.62	0.61	0.52	0.58	0.62	0.62	0.59	0.59	0.57	0.5	0.5	0.73	0.593
33	SFRP4	0.8	0.39	0.75	0.23	0.74	0.43	0.37	0.7	0.65	0.64	0.68	0.53	0.77	0.591
34	HTRA3	0.82	0.6	0.64	0.57	0.7	0.41	0.53	0.8	0.75	0.62	0.24	0.52	0.4	0.585
35	EPYC	0.77	0.7	0.73	0.41	0.57	0.23	0.24	0.79	0.65	0.79	0.24	0.74	0.72	0.583
36	PCOLCE	0.75	0.51	0.73	0.48	0.8	0.63	0.3	0.63	0.76	0.7	0.18	0.47	0.61	0.581
37	PODNL1	0.76	0.52	0.5	0.4	0.53	0.71	0.58	0.56	0.7	0.48	0.34	0.66	0.79	0.579
38	ANTXR1	0.72	0.79	0.88	0.47	0.66	0.51	0.25	0.67	0.79	0.68	0.32	0.59	0.18	0.578
39	ZNF469	0.78	0.48	0.72	0.59	0.76	0.46	0.35	0.53	0.75	0.68	0.12	0.61	0.66	0.576
40	CORIN	0.55	0.8	0.75	0.55	0.73	0.47	0.21	0.67	0.82	0.67	0.11	0.51	0.64	0.575
41	THY1	0.72	0.64	0.73	0.64	0.8	0.38	0.39	0.75	0.81	0.44	0.49	0.59	0.1	0.575
42	MFAP5	0.75	0.77	0.72	0.55	0.5	0.54	0.34	0.81	0.67	0.52	0.26	0.22	0.77	0.571
43	GLT8D2	0.7	0.63	0.76	0.53	0.74	0.55	0.51	0.76	0.82	0.7	0.27	0.31	0.08	0.566
44	ISLR	0.8	0.49	0.68	0.45	0.81	0.61	0.54	0.78	0.76	0.48	0.24	0.47	0.23	0.565
45	COL6A2	0.75	0.49	0.59	0.55	0.78	0.57	0.46	0.64	0.77	0.71	0.05	0.27	0.58	0.555
46	PDGFRB	0.7	0.53	0.71	0.5	0.68	0.39	0.56	0.47	0.77	0.77	0.36	0.48	0.21	0.548
47	LOXL2	0.6	0.68	0.62	0.5	0.57	0.56	0.27	0.64	0.73	0.7	0.32	0.59	0.3	0.545
48	OMD	0.76	0.59	0.58	0.34	0.72	0.4	0.42	0.67	0.79	0.68	0.24	0.17	0.72	0.545
49	P4HA3	0.68	0.64	0.73	0.53	0.76	0.46	0.36	0.64	0.57	0.67	0.19	0.62	0.22	0.544
50	LOX	0.74	0.71	0.76	0.19	0.47	0.34	0.36	0.68	0.58	0.72	0.14	0.7	0.68	0.544
51	DCN	0.71	0.58	0.72	0.41	0.67	0.63	0.51	0.5	0.62	0.71	0.2	0.07	0.73	0.543
52	TIMP2	0.67	0.65	0.79	0.47	0.77	0.46	0.14	0.56	0.81	0.49	0.25	0.34	0.61	0.539
53	WNT2	0.71	0.67	0.63	0.44	0.76	0.46	0.23	0.63	0.71	0.22	0.03	0.74	0.78	0.539
54	NOX4	0.64	0.78	0.77	0.54	0.76	-0	-0.2	0.79	0.67	0.41	0.57	0.71	0.58	0.538
55	TNFSF4	0.73	0.66	0.75	0.37	0.58	0.28	0.22	0.68	0.6	0.63	0.14	0.62	0.63	0.53
56	EMILIN1	0.71	0.39	0.64	0.58	0.77	0.6	0.44	0.57	0.74	0.58	0.03	0.13	0.68	0.528
57	NID2	0.7	0.73	0.57	0.64	0.76	0.39	0.37	0.7	0.81	0.57	0.16	0.49	-0	0.527
58	COL6A1	0.72	0.53	0.62	0.48	0.78	0.56	0.32	0.58	0.72	0.62	0.14	0.32	0.45	0.526
59	COL8A2	0.7	0.72	0.77	0.29	0.56	0.42	0.25	0.66	0.73	0.51	0.31	0.32	0.59	0.525
60	TNFAIP6	0.77	0.59	0.69	0.52	0.73	-0.3	0	0.75	0.62	0.72	0.34	0.67	0.68	0.522
61	RCN3	0.71	0.38	0.54	0.38	0.73	0.44	0.32	0.57	0.58	0.58	0.35	0.55	0.61	0.518
62	FAM26E	0.66	0.71	0.74	0.55	0.57	0.33	0.5	0.59	0.62	0.57	0.39	0.41	0.09	0.518
63	COL5A3	0.57	0.61	0.59	0.4	0.66	0.37	0.39	0.64	0.77	0.75	0.22	0.59	0.14	0.515
64	DACT1	0.71	0.74	0.47	0.58	0.8	0.46	0.34	0.67	0.78	0.58	0.22	0.34	-0	0.514
65	GFPT2	0.69	0.5	0.7	0.38	0.48	0.54	0.28	0.63	0.63	0.67	0.13	0.45	0.6	0.514
66	MXRA8	0.74	0.45	0.7	0.53	0.79	0.4	0.13	0.43	0.78	0.5	0.31	0.41	0.49	0.512
67	CCDC80	0.73	0.39	0.74	0.19	0.56	0.6	0.48	0.67	0.64	0.72	0.23	0.2	0.43	0.506
68	ANGPTL2	0.67	0.56	0.75	0.46	0.65	0.15	0.2	0.7	0.76	0.67	0.22	0.34	0.45	0.506
69	MMP14	0.61	0.64	0.64	0.3	0.57	0.43	-0.1	0.6	0.64	0.7	0.15	0.67	0.64	0.502
70	KIF26B	0.64	0.67	0.59	0.44	0.64	0.41	0.13	0.57	0.71	0.65	0.26	0.51	0.31	0.502
71	FIBIN	0.85	0.65	0.77	0.54	0.75	-0.1	0.01	0.61	0.77	0.53	0.46	0.31	0.29	0.499
72	MRC2	0.64	0.52	0.71	0.46	0.62	0.63	0.43	0.4	0.65	0.47	0.15	0.42	0.39	0.499
73	KCND2	0.22	0.54	0.65	0.53	0.75	0.4	0.14	0.7	0.67	0.35	0.28	0.56	0.66	0.496
74	ALPK2	0.62	0.68	0.7	0.61	0.62	-0.2	0.06	0.7	0.68	0.78	0.15	0.31	0.73	0.493
75	ADAMTS14	0.72	0.6	0.41	0.37	0.71	0.52	0.1	0.41	0.62	0.44	0.31	0.42	0.73	0.489
76	TWIST1	0.61	0.24	0.69	0.3	0.63	0.45	0.48	0.71	0.48	0.61	0.12	0.45	0.57	0.488
77	ACTA2	0.59	0.4	0.62	0.61	0.76	0.42	0.52	0.66	0.73	0.74	-0.1	0.05	0.3	0.488
78	LAMA4	0.59	0.58	0.72	0.52	0.7	0.41	0.34	0.53	0.77	0.65	0.27	0.19	0.06	0.487
79	TAGLN	0.65	0.36	0.68	0.55	0.75	0.46	0.53	0.66	0.68	0.62	-0	0.02	0.4	0.486
80	CERCAM	0.6	0.43	0.73	0.35	0.61	0.51	0.4	0.6	0.61	0.52	0.14	0.49	0.3	0.484
81	FBLN2	0.57	0.4	0.65	0.27	0.67	0.35	0.4	0.59	0.69	0.44	0.31	0.32	0.62	0.483
82	CALD1	0.67	0.53	0.81	0.53	0.67	0.32	0.07	0.55	0.79	0.61	0.07	0.15	0.49	0.482
83	ECM2	0.68	0.56	0.7	0.38	0.73	0.27	0.15	0.47	0.75	0.6	0.34	0.25	0.34	0.478
84	MEDAG	0.54	0.26	0.67	0.3	0.51	0.43	0.48	0.54	0.49	0.64	0.36	0.35	0.62	0.476
85	TMEM119	0.59	0.22	0.65	0.34	0.67	0.59	0.48	0.36	0.63	0.42	0.4	0.25	0.59	0.476
86	TGFB3	0.67	0.26	0.65	0.27	0.68	0.58	0.4	0.64	0.66	0.55	-0.2	0.36	0.64	0.475
87	EDNRA	0.55	0.56	0.57	0.44	0.52	0.34	0.52	0.39	0.71	0.71	0.06	0.26	0.54	0.475
88	FSTL1	0.56	0.59	0.83	0.39	0.66	0.3	0.06	0.56	0.61	0.58	0.1	0.35	0.57	0.474
89	C1QTNF6	0.41	0.54	0.44	0.27	0.47	0.44	0.14	0.64	0.6	0.69	0.21	0.7	0.59	0.472
90	SCARF2	0.73	0.46	0.63	0.53	0.76	0.46	0.28	0.53	0.69	0.27	0.21	0.45	0.14	0.472
91	RGS4	0.66	0.36	0.61	0.43	0.59	0.48	0.47	0.56	0.48	0.58	0.17	0.17	0.57	0.472
92	FAM180A	0.73	0.46	0.68	0.43	0.4	0.52	0.35	0.38	0.61	0.3	0.39	0.22	0.66	0.472
93	OLFML1	0.66	0.52	0.7	0.37	0.7	0.36	0.38	0.43	0.75	0.59	0.08	0.2	0.39	0.472
94	MSRB3	0.64	0.54	0.79	0.53	0.68	0.46	0.4	0.45	0.65	0.68	0.05	0.07	0.18	0.471
95	CRISPLD2	0.56	0.49	0.66	0.35	0.7	0.35	0.22	0.58	0.7	0.82	-0.1	0.16	0.6	0.47
96	GPC6	0.73	0.68	0.81	0.56	0.73	0.14	-0.1	0.61	0.71	0.41	0.27	0.29	0.24	0.469
97	SOX11	0.49	0.34	0.66	0.62	0.72	0.4	0.35	0.39	0.4	0.33	0.28	0.58	0.52	0.468
98	GPR1	0.66	0.76	0.52	0.25	0.08	0.33	0.3	0.74	0.45	0.53	0.31	0.53	0.61	0.467
99	GAS1	0.74	0.48	0.71	0.2	0.62	0.5	0.27	0.7	0.72	0.45	-0.1	0.28	0.45	0.466
100	SERPINF1	0.61	0.55	0.73	0.28	0.56	0.55	0.09	0.56	0.43	0.74	-0	0.25	0.71	0.465

Continued

	Gene Symbol	Bladder	Breast	Colorectal	Cervical	Head and Neck	Kidney Clear Cell	Kidney Papillary	Lung Adeno	Lung Squamous	Ovarian	Prostate	Stomach	Thyroid	Average Correlation
101	PODN	0.57	0.37	0.41	0.43	0.62	0.63	0.52	0.36	0.65	0.65	0.15	0.03	0.65	0.465
102	CD248	0.66	0.25	0.63	0.46	0.64	0.38	0.5	0.51	0.67	0.68	0.04	0.42	0.2	0.465
103	PLXDC1	0.52	0.36	0.5	0.36	0.59	0.42	0.44	0.3	0.67	0.73	0.48	0.31	0.36	0.465
104	MMP13	0.22	0.72	0.56	0.23	0.42	0.49	0.26	0.47	0.45	0.68	0.14	0.62	0.77	0.464
105	COL16A1	0.65	0.28	0.56	0.21	0.38	0.58	0.48	0.54	0.63	0.64	0.2	0.3	0.57	0.463
106	CPZ	0.35	0.3	0.61	0.36	0.63	0.54	0.46	0.58	0.69	0.25	0.37	0.44	0.44	0.463
107	VGLL3	0.14	0.54	0.78	0.43	0.58	0.53	0.56	0.43	0.65	0.68	-0.2	0.39	0.5	0.463
108	COL15A1	0.56	0.33	0.71	0.45	0.42	0.43	0.51	0.67	0.7	0.48	0.39	0.22	0.13	0.462
109	SGIP1	0.6	0.73	0.69	0.39	0.7	0.24	0.38	0.47	0.72	0.57	0.35	0.3	-0.2	0.461
110	CMTM3	0.37	0.52	0.77	0.42	0.6	0.48	0.14	0.54	0.41	0.48	0.17	0.54	0.54	0.46
111	LOXL1	0.33	0.46	0.55	0.26	0.61	0.61	0.17	0.54	0.58	0.63	0.32	0.29	0.63	0.46
112	NALCN	0.37	0.42	0.49	0.32	0.59	0.54	0.51	0.37	0.53	0.44	0.57	0.34	0.48	0.459
113	COL24A1	0.35	0.46	0.72	0.32	0.74	0.39	0.36	0.4	0.4	0.44	0.28	0.48	0.62	0.458
114	PLAU	0.46	0.8	0.54	0.19	0.29	0.53	0.09	0.57	0.35	0.75	0.14	0.63	0.62	0.458
115	MFAP2	0.37	0.46	0.59	0.36	0.53	0.65	0.35	0.58	0.4	0.36	0.23	0.62	0.43	0.456
116	LTBP2	0.56	0.31	0.53	0.41	0.45	0.51	0.59	0.24	0.57	0.47	0.37	0.32	0.59	0.455
117	MATN3	0.38	0.4	0.49	0.37	0.59	0.46	0.15	0.38	0.51	0.56	0.56	0.46	0.61	0.455
118	THBS1	0.63	0.52	0.44	0.45	0.34	0.2	0.32	0.54	0.56	0.71	0.16	0.37	0.67	0.455
119	CTGF	0.63	0.36	0.49	0.4	0.71	0.51	0.5	0.39	0.63	0.51	0.21	0.38	0.13	0.45
120	DPYSL3	0.59	0.65	0.7	0.38	0.39	0.52	0.61	0.54	0.38	0.47	0	0.19	0.43	0.45
121	ARSI	0.69	0.56	0.65	0.23	0.22	0.51	0.14	0.6	0.51	0.24	0.17	0.59	0.73	0.449
122	ADAMTS16	0.69	0.59	0.69	0.57	0.74	-0.1	-0.2	0.59	0.53	0.45	0.41	0.54	0.31	0.448
123	GRP	0.48	0.53	0.61	0.51	0.65	0.29	0.33	0.38	0.52	0.54	-0.1	0.31	0.72	0.447
124	RAB31	0.61	0.58	0.84	0.19	0.21	0.53	0.02	0.4	0.46	0.62	0.31	0.43	0.61	0.447
125	HTRA1	0.64	0.55	0.69	0.46	0.46	0.51	0.23	0.59	0.69	0.42	0.16	0.3	0.1	0.446
126	CLMP	0.59	0.45	0.73	0.38	0.35	0.58	0.13	0.61	0.73	0.73	0.06	0.09	0.37	0.446
127	IGFL2	0.54	0.53	0.3	0.35	0.3	0.51	0.31	0.4	0.51	0.71	0.17	0.56	0.59	0.445
128	BICC1	0.46	0.53	0.68	0.4	0.63	-0	0.04	0.5	0.84	0.67	0.15	0.35	0.54	0.443
129	PDGFRL	0.58	0.49	0.56	0.3	0.76	0.54	0.44	0.36	0.7	0.35	0.32	0.28	0.07	0.442
130	FAM101A	0.72	0.42	-0.1	0.41	0.71	0.69	0.5	0.65	0.68	0.68	-0.2	-0.2	0.7	0.442
131	SGCD	0.74	0.58	0.79	0.49	0.7	0.44	0	0.56	0.83	0.34	0.09	0.16	0.01	0.441
132	CSMD2	0.71	0.68	0.54	0.4	0.64	0.07	0.43	0.66	0.73	0.39	0.16	0.48	-0.2	0.441
133	LAMP5	0.56	0.31	0.51	0.21	0.68	0.6	0.41	0.34	0.61	0.08	0.47	0.27	0.67	0.44
134	SERPINH1	0.52	0.42	0.44	0.37	0.46	0.52	0.23	0.55	0.39	0.46	0.27	0.65	0.41	0.438
135	PDGFRA	0.48	0.38	0.41	0.33	0.53	0.58	0.51	0.58	0.6	0.48	0.17	0	0.64	0.438
136	BNC2	0.66	0.64	0.72	0.44	0.72	0.27	-0.1	0.6	0.8	0.27	0.12	0.16	0.34	0.437
137	COL4A1	0.42	0.39	0.62	0.55	0.51	0.43	0.46	0.4	0.58	0.64	0.4	0.29	-0	0.437
138	IBSP	0.45	0.42	0.5	0.43	0.56	0.42	0.44	0.5	0.28	0.28	0.16	0.6	0.61	0.435
139	ITGA5	0.59	0.52	0.71	0.42	0.34	0.54	0.41	0.59	0.47	0.69	-0.1	0.3	0.12	0.434
140	MMP16	0.45	0.36	0.55	0.51	0.69	0.45	0.41	0.42	0.62	0.25	0.2	0.3	0.43	0.434
141	ST6GALNAC5	0.72	0.13	0.8	0.54	0.49	0.47	-0.1	0.55	0.58	0.19	0.25	0.38	0.59	0.434
142	ADAM19	0.58	0.48	0.41	0.34	0.38	0.56	0.44	0.4	0.49	0.57	0.14	0.28	0.56	0.433
143	MEIS3	0.48	0.17	0.64	0.48	0.63	0.54	0.16	0.75	0.64	0.36	-0.1	0.4	0.52	0.433
144	APCDD1L	0.61	0.31	0.45	0.37	0.28	0.65	0.47	0.74	0.57	0.26	0	0.17	0.75	0.433
145	XIRP1	0.66	0.28	0.45	0.45	0.43	0.36	0.44	0.62	0.48	0.08	0.3	0.7	0.38	0.433
146	ADAMTS4	0.5	0.3	0.52	0.44	0.62	0.33	0.39	0.58	0.58	0.59	0.11	0.39	0.27	0.432
147	PDPN	0.62	0.58	0.75	0.28	0.17	0.39	0.14	0.55	0.51	0.67	-0.1	0.53	0.5	0.432
148	COLEC12	0.66	0.35	0.71	0.37	0.73	0.46	0.1	0.16	0.6	0.66	-0.1	0.07	0.74	0.428
149	MAGEL2	0.54	0.52	0.54	0.4	0.81	0.57	0.4	0.55	0.67	0.32	-0	0.4	-0.1	0.427
150	C1S	0.65	0.33	0.75	0.24	0.34	0.44	0.37	0.41	0.51	0.43	0.19	0.21	0.67	0.426
151	DDR2	0.67	0.44	0.73	0.4	0.59	0.33	0.41	0.35	0.75	0.43	0.11	0.11	0.22	0.426
152	BMP1	0.57	0.57	0.51	0.3	0.4	0.47	0.2	0.43	0.54	0.46	0.06	0.62	0.4	0.425
153	TENM3	0.67	0.56	0.71	0.38	0.35	0.29	-0	0.6	0.66	0.56	-0	0.45	0.33	0.425
154	DACT3	0.51	0.44	0.57	0.46	0.69	0.38	0.51	0.53	0.66	0.53	0.1	-0.1	0.21	0.425
155	DIO2	0.58	0.49	0.65	0.22	0.46	0.53	0.45	0.71	0.69	0.64	0.33	0.15	-0.4	0.425
156	SPOCK1	0.51	0.67	0.83	0.46	0.49	0.37	0.12	0.78	0.31	0.56	-0.2	0.5	0.08	0.425
157	GGT5	0.6	0.05	0.59	0.45	0.66	0.46	0.42	0.29	0.65	0.48	0.24	0.26	0.34	0.422
158	PALLD	0.68	0.6	0.66	0.34	0.41	0.48	0.1	0.55	0.34	0.63	0.07	0.13	0.5	0.422
159	NTNG2	0.72	0.55	0.53	0.29	0.61	0.35	0.27	0.56	0.6	0.39	-0.1	0.4	0.29	0.422
160	LGALS1	0.57	0.36	0.54	0.43	0.38	0.41	0.19	0.58	0.54	0.58	0.07	0.31	0.51	0.421
161	CPXM1	0.79	0.27	0.52	0.35	0.77	0.66	0.43	0.61	0.67	-0	-0.2	0.48	0.15	0.42
162	ADAMTS6	0.57	0.59	0.59	0.29	0.5	0.34	0.41	0.29	0.51	0.11	0.38	0.34	0.54	0.42
163	HMCN1	0.46	0.58	0.69	0.39	0.63	0.1	0.51	0.1	0.54	0.46	0.29	0.26	0.45	0.42
164	NNMT	0.68	0.32	0.67	0.29	0.61	0.19	0.23	0.17	0.5	0.56	0.15	0.34	0.75	0.42
165	PMP22	0.39	0.39	0.63	0.39	0.69	0.36	0.23	0.41	0.65	0.67	0.21	0.09	0.34	0.419
166	TGFBI	0.54	0.59	0.05	0.29	0.25	0.59	0.23	0.54	0.48	0.61	0.36	0.42	0.5	0.419
167	CLEC11A	0.71	0.36	0.48	0.35	0.77	0.49	0.19	0.58	0.54	0.16	0.15	0.47	0.19	0.418
168	CCL11	0.68	0.34	0.18	0.36	0.4	0.51	0.3	0.66	0.62	0.7	0.21	-0.1	0.59	0.418
169	ZFPM2	0.72	0.61	0.7	0.24	0.51	0.41	0.39	0.48	0.56	0.28	0.26	0.13	0.14	0.418
170	GUCA1A	0.76	0.37	0.51	0.5	0.64	0.22	0.14	0.71	0.64	0.04	0.06	0.66	0.18	0.418
171	SPOCD1	0	0.41	0.51	0.36	0.45	0.58	0.3	0.53	0.49	0.41	0.21	0.54	0.62	0.416
172	TGFB1I1	0.56	0.38	0.58	0.36	0.56	0.44	0.33	0.53	0.7	0.64	-0.1	0.08	0.32	0.416
173	KERA	0.57	0.43	0.47	0.24	0.56	0.13	0.26	0.56	0.51	0.39	0.31	0.33	0.64	0.415
174	NUAK1	0.4	0.56	0.61	0.22	0.44	0.38	0.38	0.55	0.46	0.71	0.18	0.46	0.05	0.415
175	ALDH1L2	0.68	0.59	0.4	0.33	0.39	0.52	0.37	0.61	0.37	0.39	0.14	0.27	0.34	0.415
176	RUNX2	0.24	0.6	0.47	0.15	0.57	0.51	0.26	0.53	0.49	0.64	0.04	0.36	0.52	0.414
177	CLEC5A	0.61	0.48	0.68	0.26	0.36	0.07	0.19	0.43	0.51	0.26	0.3	0.61	0.62	0.414
178	CCIN	0.61	0.54	0.42	0.43	0.6	0.48	0.22	0.54	0.55	0.09	0.12	0.32	0.45	0.413
179	MSC	0.66	0.38	0.56	0.42	0.45	0.1	0.07	0.6	0.31	0.62	0.3	0.31	0.59	0.413
180	PRR16	0.58	0.36	0.61	0.31	0.39	0.45	0.44	0.57	0.49	0.37	0.15	0.4	0.23	0.412
181	LZTS1	0.39	0.29	0.55	0.43	0.53	0.3	0.33	0.45	0.52	0.5	0.34	0.34	0.37	0.411
182	FILIP1L	0.54	0.55	0.69	0.44	0.5	0.29	0.22	0.38	0.64	0.62	0.19	0.04	0.23	0.41
183	ZNF521	0.62	0.22	0.73	0.47	0.66	0.24	0.42	0.44	0.75	0.38	0.28	0.13	-0	0.41
184	PMEPA1	0.51	0.47	0.25	0.35	0.41	0.36	0.43	0.5	0.56	0.55	-0.1	0.4	0.63	0.409
185	TMEM158	0.59	0.44	0.24	0.41	0.32	0.58	0.44	0.66	0.18	0.73	-0.3	0.35	0.67	0.409
186	CHN1	0.46	0.27	0.68	0.38	0.58	0.44	0.16	0.48	0.62	0.49	0.16	0.23	0.36	0.408
187	TSHZ3	0.48	0.55	0.67	0.24	0.44	0.57	0.47	0.63	0.64	0.48	-0	0.15	0.01	0.407
188	MICAL2	0.37	0.58	0.18	0.26	0.48	0.53	0.3	0.25	0.63	0.54	0.24	0.26	0.66	0.406
189	LMCD1	0.6	0.4	0.51	0.23	0.7	0.41	0.25	0.41	0.5	0.22	0.09	0.37	0.58	0.405
190	SNAI2	0.28	0.43	0.71	0.18	0.25	0.51	0.48	0.68	0.28	0.76	-0.2	0.45	0.45	0.405
191	GALNT15	0.55	0.46	0.63	0.55	0.68	-0.1	0.16	0.55	0.65	0.46	0.2	0.21	0.24	0.405
192	CSGALNACT2	0.47	0.43	0.64	0.23	0.42	0.23	0.33	0.37	0.52	0.39	0.37	0.31	0.55	0.405
193	BCAT1	0.52	0.64	0.76	0.31	0.42	0.45	0.24	0.12	0.31	0.08	0.19	0.62	0.58	0.403
194	AXL	0.55	0.57	0.66	0.21	0.38	0.43	0.14	0.37	0.54	0.41	0.17	0.22	0.57	0.402
195	CILP2	0.28	0.46	0.26	0.43	0.46	0.57	0.3	0.71	0.5	0.39	0.19	0.41	0.25	0.401

Table 1. *COL11A1*-correlated genes (Spearman's rank correlation) across 13 different TCGA carcinoma types. Pink denotes the top 10% *COL11A1*-correlated genes in each individual carcinoma type.

In contrast, the top 10% *COL11A1*-anticorrelated genes were not conserved across these cancer types (data not shown). Some of the top ranked *COL11A1*-anticorrelated genes in individual cancer types were associated with normal functions of these organs suggesting that they may represent normal tissue or a noninvasive tumor component. For example, the ovarian cancer top 100 *COL11A1*-anticorrelated genes present in the GSE12172 ovarian cancer dataset were primarily expressed in ovarian tumors of low malignant potential (**Fig. 4**).

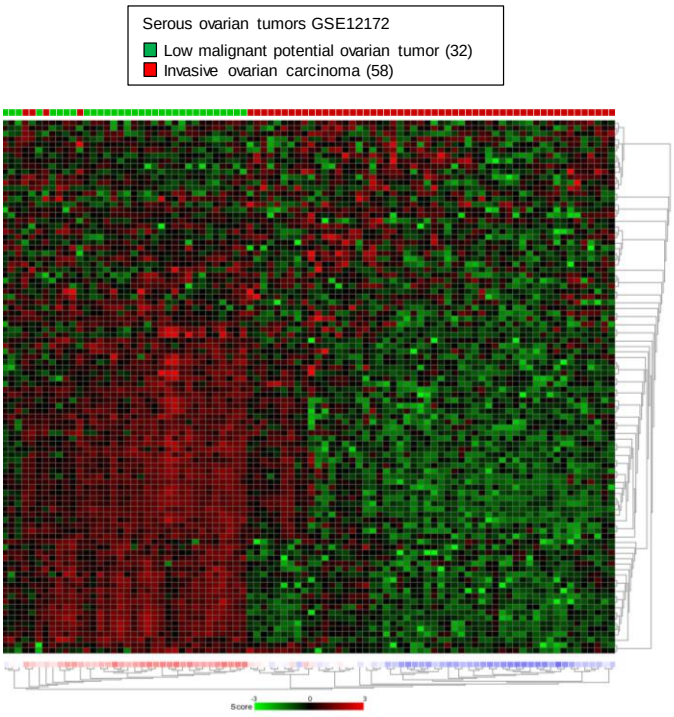

Fig. 4. Ovarian cancer *COL11A1*-anticorrelated genes are enriched in ovarian tumors of low malignant potential. GeneSet clustering analysis of low malignant potential serous ovarian tumors (n=32) and invasive serous ovarian tumors (n=58) using the top 100 *COL11A1*-anticorrelated genes (Spearman's correlation, Table S2) from the ovarian cancer TCGA dataset. The clustering analysis was performed using R2 (http://hgserver1.amc.nl/cgi-bin/r2/main.cgi).

3.3.5. Pan-cancer *COL11A1*-correlated genes are induced in cancer-activated fibroblasts

Four different public datasets were used to determine the cell type associated with expression of the pan-cancer *COL11A1*-correlated genes. Several of the 195 pan-cancer *COL11A1*-correlated genes have been shown to play a role in EMT [14]. Malignant cells undergoing EMT have been proposed as one possible source of cancer fibroblasts [15]. To determine if the pan-cancer *COL11A1*-correlated gene set is preferentially expressed in cancer cells undergoing EMT or in host-derived fibroblasts, we used the e-mtab-991 public transcription profile dataset of primary patient-derived colon cancers and their patient-derived xenografts (PDX) in nude mice [16]. Presumably, in PDX samples, fast-proliferating human cancer cells continued to grow in mice while slow-proliferating human cancer fibroblasts were lost and eventually replaced by mouse fibroblasts, which can be distinguished from human cells by species-specific gene probes [16]. GeneSet clustering analysis showed that most of the pan-cancer *COL11A1*-correlated genes had diminished levels in PDX samples in comparison to primary cancers (**Fig. 5A**), indicating that the genes are enriched in the fibroblasts rather than in the cancer cells. The preferential expression of the *COL11A1*-correlated genes in fibroblasts was further confirmed using GeneSet clustering analysis of the GSE40595 dataset in which ovarian cancer fibroblasts and epithelial cancer cells were isolated by laser-capture microdissection [17]. The pan-cancer *COL11A1*-correlated genes were preferentially expressed in cancer fibroblasts (**Fig. 5B**). In addition to

fibroblasts, immune cells are a major component of the tissue microenvironment. To exclude the possibility that the pan-cancer *COL11A1*-correlated gene set represents immune cells in the tumor microenvironment, we used the expression profile of 230 mouse hematopoietic cell types generated by the Immunological Genome Project (ImmGen) compendium [18]. In addition to immune cells, the dataset contains expression profiles of skin fibroblasts and fibroblasts residing in the thymus, lymph nodes, and spleen. The pan-cancer *COL11A1*-correlated gene set was highly represented in fibroblasts but not other hematopoietic cell lineages (**Fig. 5C**). Cancer-activated fibroblasts have a different expression profile than normal fibroblasts. Moffitt and colleagues defined a 23-gene signature of 'normal stroma' and a 25-gene signature of 'activated stroma' [19] using non-negative matrix factorization for virtual microdissection of primary and metastatic pancreatic ductal cancer samples into cell subsets with prognostic and biologic relevance. None of the 23 (0%) 'normal stroma' genes in contrast to 18 of 25 (72%) 'activated stroma' genes were present in the *COL11A1*-correlated gene set, respectively (**Fig. 5D**), suggesting that the *COL11A1*-correlated gene set represents cancer-activated fibroblasts. To determine whether the pan-cancer *COL11A1*-correlated gene set is associated with patient survival in the ~18,000 cases of liquid and solid malignancies in the PRECOG dataset [20], we compared survival z-scores for the 195 pan-cancer *COL11A1*-correlated genes with the survival z-scores for all genes in the dataset. This analysis showed that expression of the pan-cancer *COL11A1*-correlated gene set is significantly associated with poor survival (**Fig. 5E**).

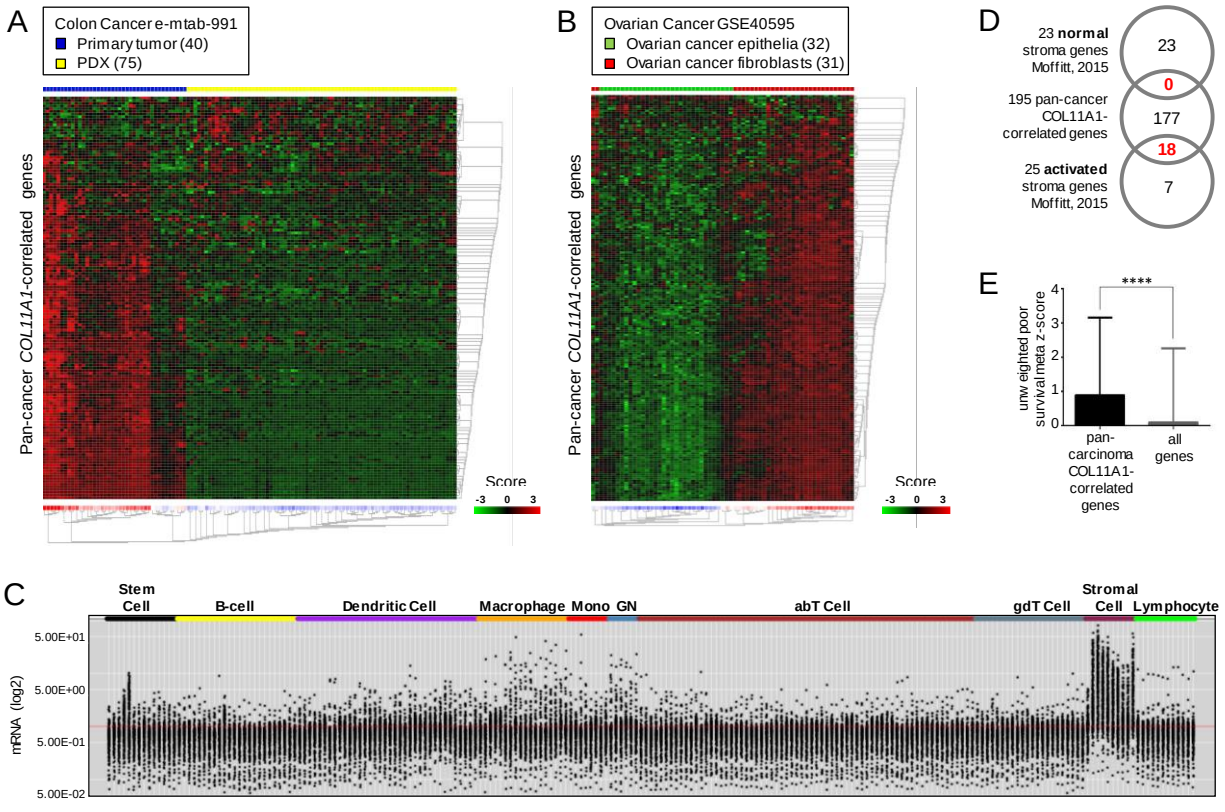

Fig. 5. The pan-cancer *COL11A1*-correlated gene set is expressed in cancer-activated fibroblasts and associated with poor patient survival in multiple cancer types. (A) GeneSet expression clustering analysis of 40 primary colon cancer samples and 75 patient-derived

xenograft (PDX) samples in the e-mtab-991 dataset using the pan-cancer *COL11A1*-correlated genes. **(B)** GeneSet expression clustering analysis of laser-microdissected ovarian cancer epithelial cells (32 samples) and fibroblasts (31 samples) in high grade serous ovarian cancer in the GSE40595 dataset using the pan-cancer *COL11A1*-correlated genes. The Euclidean distance clustering analysis heatmaps in (A) and (B) were generated using the public R2 GeneSet Clustering Analysis tool (hgserver1.amc.nl/cgi-bin/r2/main.cgi). **(C)** Expression of the pan-cancer *COL11A1*-correlated genes mapped on the transcriptome of individual murine hematopoietic and stromal cell types in the ImmGene project (immgen.com). The plot was generated using MyGeneset tool (rstats.immgen.org/MyGeneSet). Black dots represent mRNA levels (y axis) of 186 pan-cancer *COL11A1*-correlated genes (9 genes were not present in the database) across 230 individual cell types (X axis) grouped into 10 main groups. **(D)** Overlap of the 195 pan-cancer *COL11A1*-correlated genes with 23 'normal stroma' and 25 'activated stroma' genes defined by Moffitt et al. **(E)** Unweighted meta z-scores of 191 *COL11A1*-correlated genes (4 genes were not available in the PRECOG database) compared with those of all genes in the PRECOG database. The plot was generated using GraphPad Prism software version 6.0. Intergroup differences were assessed by the Student's t-test. Mean ± SEM of pan-cancer *COL11A1*-correlated genes (0.8916 ± 0.1641 N=191); Mean ± SEM of all genes (0.09918 ± 0.01416 N=23287); ****p<0.0001.

References

1. Ozdemir, B.C., et al., *Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival.* Cancer Cell, 2014. 25(6): p. 719-34.
2. Laklai, H., et al., *Genotype tunes pancreatic ductal adenocarcinoma tissue tension to induce matricellular fibrosis and tumor progression.* Nat Med, 2016. 22(5): p. 497-505.
3. Quail, D.F. and J.A. Joyce, *Microenvironmental regulation of tumor progression and metastasis.* Nat Med, 2013. 19(11): p. 1423-37.
4. Kalluri, R. and M. Zeisberg, *Fibroblasts in cancer.* Nat Rev Cancer, 2006. 6(5): p. 392-401.
5. Junttila, M.R. and F.J. de Sauvage, *Influence of tumour micro-environment heterogeneity on therapeutic response.* Nature, 2013. 501(7467): p. 346-54.
6. Beach, J.A., et al., *Sphingosine kinase 1 is required for TGF-beta mediated fibroblast-to-myofibroblast differentiation in ovarian cancer.* Oncotarget, 2016. 7(4): p. 4167-82.
7. Cheon, D.J., et al., *A collagen-remodeling gene signature regulated by TGF-beta signaling is associated with metastasis and poor survival in serous ovarian cancer.* Clin Cancer Res, 2014. 20(3): p. 711-23.
8. Vazquez-Villa, F., et al., *COL11A1/(pro)collagen 11A1 expression is a remarkable biomarker of human invasive carcinoma-associated stromal cells and carcinoma progression.* Tumour Biol, 2015. 36(4): p. 2213-22.
9. Wu, Y.H., et al., *COL11A1 promotes tumor progression and predicts poor clinical outcome in ovarian cancer.* Oncogene, 2014. 33(26): p. 3432-40.
10. Tan, T.Z., et al., *CSIOVDB: a microarray gene expression database of epithelial ovarian cancer subtype.* Oncotarget, 2015. 6(41): p. 43843-43852.
11. Desmouliere, A., C. Guyot, and G. Gabbiani, *The stroma reaction myofibroblast: a key player in the control of tumor cell behavior.* Int J Dev Biol, 2004. 48(5-6): p. 509-17.

12. Lawrenson, K., et al., *NPPB is a novel candidate biomarker expressed by cancer-associated fibroblasts in epithelial ovarian cancer.* Int J Cancer, 2015. 136(6): p. 1390-401.
13. Shin, G., et al., *GENT: gene expression database of normal and tumor tissues.* Cancer Inform, 2011. 10: p. 149-57.
14. Tan, T.Z., et al., *Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients.* EMBO Mol Med, 2014. 6(10): p. 1279-93.
15. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* Journal of Clinical Investigation, 2009. 119(6): p. 1420-1428.
16. Julien, S., et al., *Characterization of a large panel of patient-derived tumor xenografts representing the clinical heterogeneity of human colorectal cancer.* Clin Cancer Res, 2012. 18(19): p. 5314-28.
17. Yeung, T.L., et al., *TGF-beta modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment.* Cancer Res, 2013. 73(16): p. 5016-28.
18. Shay, T. and J. Kang, *Immunological Genome Project and systems immunology.* Trends Immunol, 2013. 34(12): p. 602-9.
19. Moffitt, R.A., et al., *Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma.* Nat Genet, 2015. 47(10): p. 1168-78.
20. Gentles, A.J., et al., *Integrating Tumor and Stromal Gene Expression Signatures With Clinical Indices for Survival Stratification of Early-Stage Non-Small Cell Lung Cancer.* J Natl Cancer Inst, 2015. 107(10).

4) other achievements

- **What opportunities for training and professional development has the project provided?**

Nothing to report.

- **How were the results disseminated to communities of interest?**

Nothing to report.

- **What do you plan to do during the next reporting period to accomplish the goals?**

During the no-cost extension of the project, we will validate the gene signatures in patient samples and develop a preliminary quantitative assay for use in the clinical setting. The development of a reliable test for the identification of high risk patients is not only crucial to improving their clinical management but also timely because of the emergence of personalized treatment strategies for ovarian cancer. A validated gene signature to identify patients with adverse outcomes has the potential to reduce both the human and financial costs of ineffective therapies and associated toxicities.

4. **IMPACT:**

- **What was the impact on the development of the principal discipline(s) of the project?**

Future efforts to specifically target activated cancer fibroblasts in ovarian cancer can be improved by designing novel therapies that target genes such as *COL11A1*, which are specifically expressed in activated cancer fibroblasts. This provides an opportunity to deliver targeted therapies directed at the underlying mechanism of the poor prognosis signature. This will facilitate more individualized treatment decisions and improve the quality of care for patients with ovarian cancer. In contrast to αSMA and other markers frequently used in mouse models, COL11A1 *is not* expressed in mesenchymal precursors in normal organs or fibroblasts associated with non-cancerous conditions, such as inflammation and organ fibrosis. This should increase targeting efficacy and reduce off-target toxicity.

- **What was the impact on other disciplines?**

The results of our analysis have an important therapeutic implication for solid tumors other than ovarian. The strong association of the COL11A1-coexpressed gene signature with poor survival across different cancer types suggests that therapeutic approaches targeting fibroblast activation in ovarian cancer should be effective in multiple cancers.

- **What was the impact on technology transfer?**

Nothing to report.

- **What was the impact on society beyond science and technology?**

Nothing to report.

5. **CHANGES/PROBLEMS:**

- **Changes in approach and reasons for change**. **No.**
- **Actual or anticipated problems or delays and actions or plans to resolve them.**

Reasons for 6-month no-cost extension have been previously reported and approved.

- **Changes that had a significant impact on expenditures**. **No.**
- **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents. No.**
- **Significant changes in use or care of human subjects. No.**
- **Significant changes in use or care of vertebrate animals. No.**
- **Significant changes in use of biohazards and/or select agents. No**

6. **PRODUCTS:**

Nothing to report.

- **Publications, conference papers, and presentations**
  - **Journal publications.**

Liu Z, Beach JA, Agadjanian H, Jia D, Aspuria P-A, Karlan BY, **Orsulic S**. Suboptimal cytoreduction in ovarian carcinoma is associated with molecular pathways characteristic of increased stromal activation. Gynecologic Oncology 2014; 139:394-400. *Lead Article with Editorial Review in Gynecologic Oncology; Editorial Review in Obstetrical & Gynecological Survey.* (Published, acknowledged grant funding)

Beach JA, Aspuria P-J, Cheon D-J, Agadjanian H, Walsh CS, Karlan BY, **Orsulic S**. Sphingosine kinase 1 is required for TGF-β mediated fibroblast-to-myofibroblast differentiation in ovarian cancer Oncotarget 2015; 7:4167-4182. (Published, acknowledged grant funding)

  - **Books or other non-periodical, one-time publications. N/A**
  - **Other publications, conference papers, and presentations. N/A**
- **Website(s) or other Internet site(s). N/A**
- **Technologies or techniques. N/A**
- **Inventions, patent applications, and/or licenses. N/A**
- **Other Products. N/A**

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

- **What individuals have worked on the project?**

Name:	Sandra Orsulic
Project Role:	PI
Nearest person month worked:	1.2
Contribution to Project:	Dr. Orsulic oversaw statistical analyses of the gene signatures, analyzed and interpreted the data, wrote and published one manuscript, and prepared another one for publication.

Name:	Dongyu Jia
Project Role:	Postdoctoral fellow
Nearest person month worked:	3
Contribution to Project:	Dr. Jia organized retrieval of pathology samples, isolated RNA and prepared cDNA from samples, optimized quality control for Nanostring assay, and assisted in the writing of the published manuscript.

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

Nothing to report.

- **What other organizations were involved as partners?**

Nothing to report.

8. **SPECIAL REPORTING REQUIREMENTS**

- **COLLABORATIVE AWARDS:** N/A
- **QUAD CHARTS:** N/A

9. **APPENDICES:** N/A